

РУКОВОДСТВО SMART

Часть 1

НАЧИНАЯ РАБОТАТЬ С XML

Начала работы с форматом XML

XML является полезным инструментом для хранения и передачи данных и может служить дополнительным первым шагом на пути к внедрению документов из серии SMART стандартов. XML был разработан для прочтения его как человеком, так и машиной, что делает его идеальной технологией для построения процесса публикации.

Как язык разметки, XML может использоваться для маркировки, категоризации и структурной организации содержимого стандартов, что открывает новые возможности для публикации, отображения и использования стандартов в Интернете.

ИСО выпускает XML-версии стандартов, которые все чаще используются национальными органами по стандартизации для предоставления пользователям расширенных цифровых услуг. К ним относятся базы данных, получаемые по подписке, и услуги по извлечению необходимых требований.

Недавно запущенная платформа ИСО/МЭК Online Standards Development (OSD) предоставляет техническому сообществу уникальное и гармонизированное пространство для совместной работы. Являясь инструментом для публикации документов на основе XML, платформа OSD закладывает основу для создания новых продуктов и услуг с высокой добавленной стоимостью для мира будущего.

Данное руководство поможет вам начать работу с XML и расскажет о дополнительных источниках информации.

Обзор XML

- XML расшифровывается как eXtensible Markup Language (расширяемый язык разметки)
- XML был разработан для хранения и передачи данных
- XML был разработан как для человеческого, так и для машинного чтения
- XML часто используется для распространения данных через сеть Интернет
- В XML каждый фрагмент данных помечен тегами, которые идентифицируют его
- Используя XML, организации могут устанавливать свои собственные теги для обмена информацией

Что такое XML?

XML - это сокращение от eXtensible Mark-up Language. Это язык тегов, используемый для обозначения типов текста для использования в машинном контексте. Например, он указывает компьютеру или приложению, является ли определенная строка текста в документе заголовком или абзацем. Если вы знакомы с HTML, то XML - это аналогичная концепция, но вместо того, чтобы размечать содержимое для использования в Интернете, XML размечает содержимое документов для широкого спектра целей.

XML существует в различных «типах» или «вкусах»; такие «типы» XML называются схемами. Схема выбирается в зависимости от типа маркируемого документа. Например, журнальные статьи часто маркируются с помощью схемы JATS, пьесы могут маркироваться с помощью TEI, а учебники - с помощью DocBook. ИСО, МЭК и СЕН используют схему под названием NISO STS. Руководство, разработанное ИСО во взаимодействии с МЭК по совместному использованию NISO STS, можно найти на сайте ИСО [здесь](#) и [здесь](#).

Рассматривайте XML как набор правил, определяющих, как должны быть расположены данные. Теги используются для обозначения элементов данных, подобно тому, как используются скобки в математических уравнениях. С помощью данных тегов можно идентифицировать и описать содержащуюся в них информацию.

The diagram illustrates the relationship between XML code and a technical report cover. On the left, the cover of 'TECHNICAL REPORT ISO/TR 19686-2' is shown, including the title 'Petroleum products — Equivalency of test method determining the same property — Part 2: Density of petroleum products' and its French equivalent. On the right, the corresponding XML code is displayed, with arrows pointing from the code to the cover elements. The XML code uses tags like `<title-wrap xml:lang="en">` and `<title-wrap xml:lang="fr">` to wrap the title in English and French, and `<intro>`, `<main>`, `<compl>`, and `<full>` to structure the introductory text, main content, completion, and full title respectively.

XML против HTML

XML и HTML создавались с разными целями:

- XML был разработан для передачи данных - с акцентом на то, что представляют собой данные
- HTML был разработан для отображения данных - с акцентом на то, как данные выглядят
- Теги XML не предопределены, так, как теги HTML
- XML часто является дополнением к HTML

XML обеспечивает семантическую фиксацию текста в цифровой среде. Другими словами, он придает тексту смысл в цифровом контексте. Это позволяет:

- извлекать информацию из текста (например, термины и определения стандартов ИСО доступны для поиска и отображения на платформе Online Browsing Platform)
- преобразовывать документ в другие форматы и другие компьютерные языки
- автоматизировать работу с документами в рамках издательского процесса
- получать информацию в последовательном виде (идеально для документов с предопределенной структурой, таких как стандарты)

Каковы преимущества работы с XML?

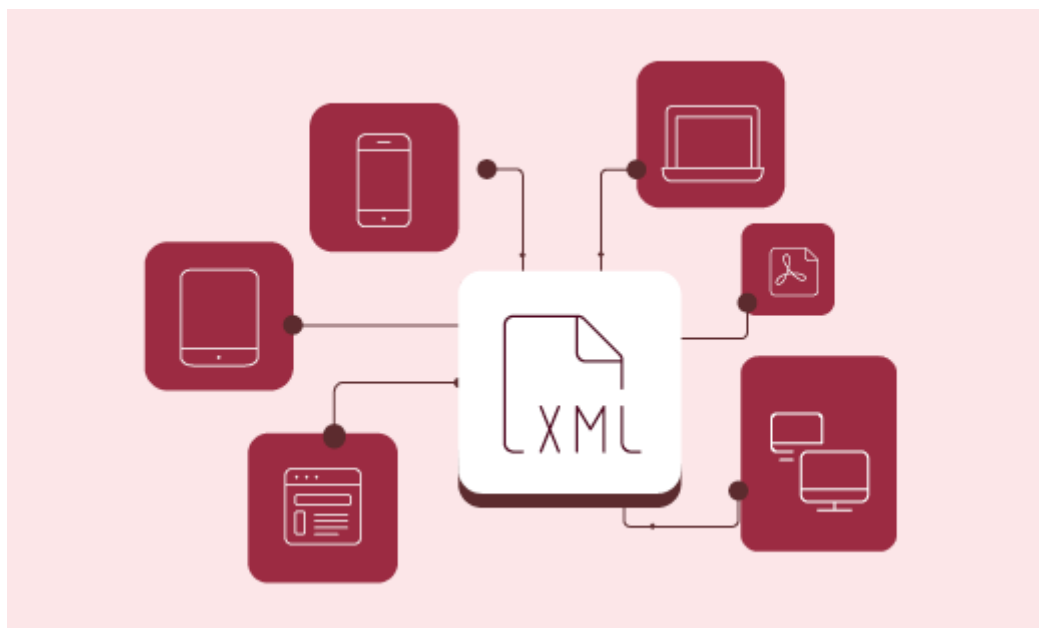
XML позволяет национальным органам по стандартизации разрабатывать и продавать дополнительные услуги для поддержки пользователей стандартов, такие как:

- Динамическое краткое представление
- Полнотекстовый поиск с повышенной точностью результатов
- Богатая навигация внутри документа (внутренние ссылки) и по внешним документам (внешние ссылки на все стандарты или нормативную документацию)
- Дополнительные функции (например, отображение изменений, подчеркивание требований, терминов, определенных в стандарте)
- Динамическое дополнение контента (например, примечания внутри текста или рисунков)
- Активация формул и таблиц
- Самостоятельная оценка соответствия содержимого требованиям стандартов

XML часто используется для публикаций и обмена структурированным содержимым.

Гибкость XML позволяет публиковать контент в различных форматах, включая:

- Базы данных
- Приложения
- Веб-страницы
- PDF
- EPUB



XML в ИСО

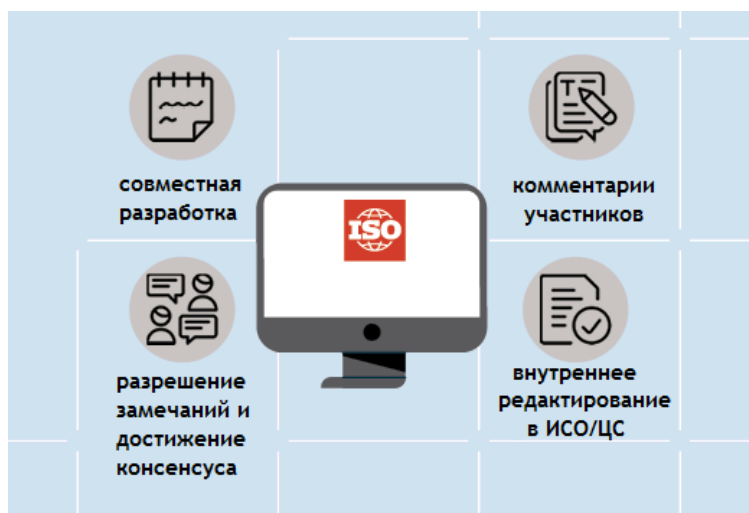
ИСО использует различные инструменты (в том числе Fonto и eXstyles) для создания XML-версий стандартов. XML применяется в различных целях, в том числе:

- Для создания окончательной версии стандарта в формате PDF. Порядок преобразования в производственной цепочке обычно следующий: Word → XML → PDF. Большинство документов проходит через «XML обработку». Остальные проходят через Word → PDF, при этом продавец выпускает XML-версию после публикации документа («процесс бэк-каталога»).
- Для наполнения Online Browsing Platform (OBP). OBP может отображать HTML-версию (XML→ HTML) или PDF, если XML-версия отсутствует. Если стандарт имеет XML-версию, то поиск терминов и определений может осуществляться с помощью функции поиска в OBP.
- Для разработки новых услуг для национальных органов по стандартизации. Все большее число национальных органов используют XML-версии стандартов для создания новых продуктов и услуг для своих локальных рынков.

Онлайн-разработка стандартов (Online Standards Development (OSD))

Новая платформа Online Standards Development (OSD) предоставляет техническому сообществу возможность совместной разработки стандарта, начиная с предварительной стадии и заканчивая публикацией.

OSD представляет собой инструмент разработки, основанный на XML. Данная программа полностью исключает Word, то есть, нет необходимости преобразовывать документ в XML - он создан уже в XML. Инструмент использует ту же XML-схему, которую применяют ИСО и МЭК для своих XML-документов, NISO STS.



Являясь средством публикации на основе XML, платформа OSD помогает заложить основу для создания новых продуктов с добавленной стоимостью:

- Сотрудничество: Совместная работа экспертов над стандартом в режиме реального времени
- Точность и качество: Платформа берет на себя заботу о структуре и оформлении документа, чтобы эксперты могли сосредоточиться на содержании
- Простота: комментирование и обработка замечаний в режиме онлайн
- Интеграция: Платформа доступна через существующие средства разработки стандартов ИСО и МЭК
- Гармонизированный подход: Единый процесс разработки стандартов МЭК/ИСО
- NISO STS: Система разработки контента, основанная на признанных рынком стандартах

XML в национальных органах по стандартизации

Около 20 национальных органов используют XML в своих цепочках производства стандартов, а также для создания новых продуктов и услуг для своих локальных рынков. Примерами расширенных услуг, предлагаемых некоторыми из них, являются:

AFNOR

Подборка услуг Французской ассоциации по стандартизации (AFNOR):

- **COBAZ** - комплексное решение для идентификации, консультирования, мониторинга и разбора стандартов.
- **Нахождение требований** - помогает определить (внутри нормативного текста) обязательные положения, которые должны быть выполнены, и положения, которые полезно знать (например, то, что допускается и рекомендуется)

DIN

Подборка услуг Немецкого института стандартизации (DIN)

- **Nautos** - справочная база данных, помогающая осуществлять поиск и управление стандартами и техническими правилами.
- **Beuth e-Norm** - браузерное программное решение для управления документами и их обновлениями.
- **Beuth Standards Manager** - аутсорсинговая услуга, включающая задачи, связанные с закупками, управлением лицензиями и администрированием документной платформы.
- **DITR Data service** - обеспечивает централизованный мониторинг всех новых, измененных и отозванных документов.

SN

Standards Norway (SN) предлагает услугу **«Extended Standards» («Расширенные стандарты»)**, которая предоставляет стандарт в виде открытого файла, что облегчает извлечение и использование необходимой информации, включая управление требованиями.

SN также имеет **лицензионные соглашения** с разработчиками программного обеспечения для предоставления цифровых услуг, включая Corporater, которая предоставляет интегрированную GRC-платформу для управления рисками и соблюдением нормативных требований.

Контрольный список для начала работы

Работа с XML может помочь преобразовать ваш контент в данные и позволить создавать новые продукты и услуги. К числу ключевых моментов относятся:

- **Обратная связь от пользователей:** С помощью опросов и семинаров выясните, какой вид цифрового контента/услуг поможет вашим пользователям в удовлетворении их конкретных потребностей (например, создание базы данных с возможностью быстрого поиска). Возможно, стоит разделить пользователей на несколько широких категорий (например, малые и средние предприятия и крупные корпорации).
- **Оценка контентной стратегии:** Проанализируйте текущий рабочий процесс создания контента и определите проблемные моменты. Как осуществляется совместная работа авторов и какие инструменты они используют (например, Word)? Есть ли ощутимые преимущества от привлечения внешних продавцов или разработки/лицензирования новых инструментов?
- **Преобразование в формат XML:** Подумайте, хотите ли вы перевести существующие документы (например, каталог стандартов) в формат XML. В первую очередь необходимо ответить на следующие вопросы:
 - Какое содержимое должно быть сохранено в существующих документах?
 - Сколько у вас имеется документов (и их типов)? Согласовано ли их форматирование?
 - Каков наилучший формат предоставления для каждого типа содержимого?
 - Как должна быть организована система ссылок?
- **Услуги по преобразованию контента:** Возможно, вы захотите передать преобразование своего каталога в XML на аутсорсинг, для чего существует множество специализированных поставщиков. Цена (обычно за страницу или документ) будет зависеть от ряда факторов, включая:
 - Какой исходный формат материала вы конвертируете (например, Word и/или PDF?)
 - Какой тип документа вы конвертируете?
 - Требуется ли для конвертации экспертная оценка содержания?
 - Требуется ли конвертирование графики или повторное создание контента?
 - Будут ли полезными средства автоматизированного преобразования?
- **Средства авторинга/публикации:** Готова ли ваша команда перейти к использованию средств публикации на основе XML для создания нового контента (например, по аналогии с новой платформой ICO OSD)? Существует множество специализированных поставщиков, у которых можно лицензировать соответствующий инструментарий, но при этом необходимо учитывать требования к обучению и потенциальное неприятие изменений.
- **Разработчики программного обеспечения:** Возможно, вы захотите передать лицензию на XML-контент разработчику программного обеспечения для создания (и, возможно, распространения) новых услуг для ваших конечных пользователей. Перед заключением любого коммерческого соглашения необходимо получить юридическую консультацию.

Доступная поддержка ИСО:

Семинары: ИСО проводит семинары по работе с форматом XML

Консультации: Можно организовать встречу со штатными экспертами центрального секретариата ИСО по вопросам работы с XML

Оценка цифровой зрелости: Примите участие в исследовании Digital Maturity Assessment, которое поможет вам определить следующие шаги на пути цифровой трансформации

OSD: Используйте платформу Online Standards Development (OSD) и ознакомьтесь с подходом к созданию структурированного контента/данных на основе XML

Вопрос-ответ

Какие национальные органы по стандартизации в настоящее время работают с XML-версиями стандартов?

В настоящее время XML используют примерно 20 национальных органов по стандартизации, в том числе: AFNOR, BSI, DKE, DIN, SA, SFS, SIS, SN и UNI.

Нужно ли мне работать с XML, чтобы использовать SMART стандарты?

Не обязательно, но XML - это эффективный способ организации контента и преобразования информации в данные.

Какие форматы хранения контента альтернативны XML?

JSON, HTML и YAML.

В чем разница между HTML и XML?

XML был разработан для передачи данных, при этом основное внимание уделялось тому, что они собой представляют. HTML был разработан для отображения данных с упором на то, как они выглядят. Теги XML не являются предустановленными, как теги HTML.

Нельзя ли просто преобразовать существующие документы Word или PDF в XML?

Можно. Несмотря на то, что процесс публикации на основе XML имеет свои преимущества, существуют специализированные организации, которым можно передать свои документы в формате Word или PDF для преобразования в XML.

Разве PDF-документы не являются машиночитаемыми?

С технической точки зрения PDF-файл может содержать метаданные, код и язык, которые могут быть распакованы и обработаны машиной. Однако он считается форматом, содержащим много изображений и не предназначенным для передачи данных.

Какие форматы являются машиночитаемыми?

XML, HTML, онтологические языки (например, OWL, ReqIF, или RDF), JSON, SQL, и QR-коды.

Дополнительные ресурсы

Материалы ИСО

- [Программа ИСО SMART](#)
- [Руководство ИСО/МЭК по NISO STS](#)
- [Вспомогательные материалы NISO по комплексу тегов стандартов \(STS\)](#)
- [Учебные материалы по пользованию платформой OSD](#)

Внешние ресурсы, посвященные формату XML

- [Введение в XML \(w3schools.com\)](#)
- [Краткий курс по XML - XML | Краткий курс | Пошаговое руководство для начинающих | Udemy](#)
- [Основной курс по XML \(linkedin.com\)](#)
- [XML: Основы: Все наименования: Изучите данные темы: UITS IT Training: Университет Индианы \(iu.edu\)](#)